

A Multiway Relay Channel with Balanced Sources

Lawrence Ong and Roy Timo

Abstract—We consider a joint source-channel coding problem on a finite-field multiway relay channel, and we give closed-form lower and upper bounds on the optimal source-channel rate. These bounds are shown to be tight for all discrete memoryless sources in a certain class \mathcal{P}^* , and we demonstrate that strict source-channel separation is optimal within this class. We show how to test whether a given source belongs to \mathcal{P}^* , we give a balanced-information regularity condition for \mathcal{P}^* , and we express \mathcal{P}^* in terms of conditional multiple-mutual informations. Finally, we show that \mathcal{P}^* is useful for a centralised storage problem.

I. INTRODUCTION

THE *multiway relay channel* is a multicast network model in which many users exchange data via a relay [1, 2]. The model is widely applicable to wireless cellular [3], satellite [4], mesh [5] networks, and storage networks [6], and its information-theoretic limits will provide design insights for future cooperative communications systems. Despite much recent attention [2, 7]–[14], the channel’s information-theoretic limits remain largely unknown. In this work, we consider the limits of the following setup:

- $L \geq 2$ users have correlated data that need to be exchanged via the relay. The correlated data are generated by an arbitrary discrete memoryless source.
- The uplink channel (users to relay) and the downlink channel (relay to users) are memoryless additive-noise channels defined over an arbitrary finite field.

The discrete memoryless source serves as a simple model for distributed correlated data in, for example, cloud storage systems, sensor networks, and mobile applications [15]–[18]. The *finite-field channel* both generalises the binary-symmetric channel and serves as a stepping stone to other important linear additive-noise channels, such as the Gaussian multiway relay channel.

An efficient communications system for the above problem needs to effectively integrate distributed data compression with multiuser channel coding. For such systems, an important information-theoretic benchmark is the *optimal source-channel rate*—the minimum number of channel uses per source symbol needed for reliable communications. The main problem of interest in this paper is to determine the optimal source-channel rate.

Ong et al. [19] studied a limited version of the above problem with three users. They determined the optimal source-channel rate for sources with specific entropic structures, and demonstrated that strict source-channel separation is optimal for such class of sources. The present paper strengthens and

generalises the main ideas and results of Ong et al. [19] to three or more users.

In Section II, we present lower and upper bounds on the optimal source-channel rate that hold for any source and $L \geq 2$ users. The upper bound (i.e., achievability) is proved using a standalone distributed source code proposed by Timo et al. [7] together with a standalone *functional-decode-forward* channel code by Ong et al. [2].

We show in Section III that the above lower and upper bounds coincide for a class of sources \mathcal{P}^* —regardless of the channel parameters—and the result is a closed-form expression for the optimal source-channel rate. The class \mathcal{P}^* is computable in the usual information-theoretic sense, and it is determined by the underlying distributed source-coding problem. We show how to test whether or not any given source belongs to \mathcal{P}^* by solving a certain linear system.

In Section IV, we give a *balanced information* regularity condition for \mathcal{P}^* that can be used whenever the methods in Section III are either impractical or undesirable. The balanced-information condition is expressed in term of *conditional multiple-mutual informations* [20]–[22], which can be visualised using information diagrams and the *I*-measure formalism of Yeung [23]. We use this approach to determine the optimal source-channel rate of some sources.

Finally, in Section V, we conclude the paper by considering a centralised storage problem with L clients. The class of sources \mathcal{P}^* plays an important role in this problem, and we show how the results of Sections III and IV can be used to describe the optimal storage rate.

II. OPTIMAL SOURCE-CHANNEL RATE

A. Notation

We denote random variables by uppercase letters, e.g. W ; their alphabets by matching calligraphic font, e.g. \mathcal{W} ; and elements of an alphabet by lowercase letters, e.g. $w \in \mathcal{W}$. The Cartesian product of \mathcal{W} and \mathcal{W}' is $\mathcal{W} \times \mathcal{W}'$, and the m -fold Cartesian product of \mathcal{W} is \mathcal{W}^m . For integers a and b , with $a \leq b$, we let $[a, b] := \{a, a+1, \dots, b\}$. Subsets and strict subsets are identified by \subseteq and \subset respectively. We will often consider subsets $\mathcal{S} \subseteq [1, L]$ and, in such cases, we let $\mathcal{S}^c := [1, L] \setminus \mathcal{S}$ denote the complement of \mathcal{S} . When \mathcal{S} is a singleton $\{\ell\}$ or the complement of a singleton $\{\ell\}^c$, we write $\ell = \{\ell\}$ and $\ell^c = \{\ell\}^c$. We let $\|\mathbf{r}\| := |r_1| + |r_2| + \dots + |r_L|$ denote the L^1 norm of a real-valued vector $\mathbf{r} \in \mathbb{R}^L$. The base of all logarithms in this paper is two.

B. Source model

Consider L arbitrarily-dependent discrete random variables

$$(W_1, W_2, \dots, W_L), \quad (1)$$

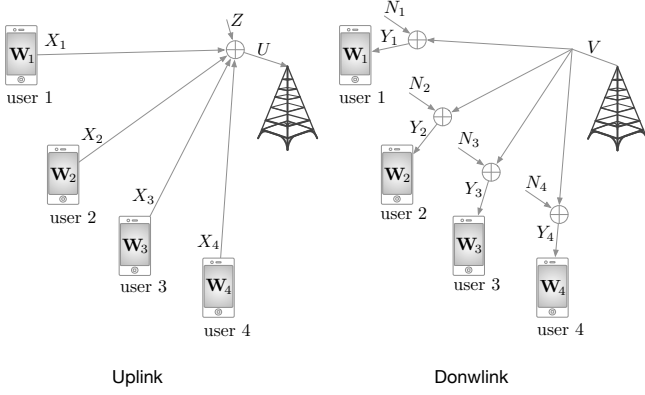


Fig. 1. The uplink and downlink channel laws of the finite-field multiway relay channel with four users.

where the ℓ -th variable W_ℓ is defined on an alphabet \mathcal{W}_ℓ and associated with user ℓ . Let

$$(\mathbf{W}_1, \mathbf{W}_2, \dots, \mathbf{W}_L) := \{(W_{1,t}, W_{2,t}, \dots, W_{L,t})\}_{t=1}^m$$

be a string of m independent and identically distributed (iid) copies of (1) indexed by t . The source data of user ℓ is the iid m -tuple $\mathbf{W}_\ell = (W_{\ell,1}, W_{\ell,2}, \dots, W_{\ell,m})$. Each user is required to exchange its source data with that of every other user.

C. Channel model

The uplink channel (users to relay) and downlink channel (relay to users) are both memoryless and defined over a finite field \mathcal{F} equipped with addition \oplus . The per-symbol law characterising the memoryless uplink channel is

$$U := X_1 \oplus X_2 \oplus \dots \oplus X_L \oplus Z, \quad (2a)$$

where $X_\ell \in \mathcal{F}$ is the symbol sent by user ℓ , U is the symbol observed by the relay, and $Z \in \mathcal{F}$ is independent arbitrarily-distributed additive noise. Similarly, the memoryless downlink is

$$Y_\ell := V \oplus N_\ell, \quad \ell \in [1, L], \quad (2b)$$

where $V \in \mathcal{F}$ is sent by the relay, $Y_\ell \in \mathcal{F}$ is observed by user ℓ , and $N_\ell \in \mathcal{F}$ is independent additive noise at user ℓ 's receiver. Figure 1 depicts the setup for $L = 4$ users. The uplink and downlink are memoryless; that means Z and all N_ℓ 's are independent, and they are each iid over all channel uses.

D. Codes for full data exchange

An (m, n) -code for exchanging the users' source data is specified by a collection of mappings

$$\{f_{1,t}, \dots, f_{L,t}, \phi_t\}_{t=1}^n \quad \text{and} \quad \{g_1, \dots, g_L\}, \quad (3)$$

where $f_{\ell,t} : \mathcal{W}_\ell^m \times \mathcal{F}^{t-1} \rightarrow \mathcal{F}$, $\phi_t : \mathcal{F}^{t-1} \rightarrow \mathcal{F}$ and $g_\ell : \mathcal{W}_\ell^m \times \mathcal{F}^n \rightarrow \mathcal{W}_1^m \times \mathcal{W}_2^m \times \dots \times \mathcal{W}_L^m$. We assume that the nodes operate in the full-duplex mode, and the uplink and the downlink are perfectly synchronised. During the t -th channel use, each user $\ell \in [1, L]$ sends

$$X_{\ell,t} := f_{\ell,t}(\mathbf{W}_\ell, Y_{\ell,1}, \dots, Y_{\ell,t-1})$$

over the uplink, and the relay sends

$$V_t := \phi_t(U_1, U_2, \dots, U_{t-1})$$

over the downlink. User ℓ observes $Y_{\ell,t}$ as per (2b) and the relay observes U_t as per (2a). After n channel uses, user ℓ has observed n symbols $\mathbf{Y}_\ell = (Y_{\ell,1}, \dots, Y_{\ell,n})$ from the downlink. It outputs

$$(\hat{\mathbf{W}}_{\ell,1}, \hat{\mathbf{W}}_{\ell,2}, \dots, \hat{\mathbf{W}}_{\ell,L}) := g_\ell(\mathbf{W}_\ell, \mathbf{Y}_\ell),$$

where $\hat{\mathbf{W}}_{\ell,i}$ denotes its reconstruction of \mathbf{W}_i . Let

$$P_e := \mathbb{P} \left[\bigcup_{\ell=1}^L \left\{ g_\ell(\mathbf{W}_\ell, \mathbf{Y}_\ell) \neq (\mathbf{W}_1, \mathbf{W}_2, \dots, \mathbf{W}_L) \right\} \right]$$

denote the average probability of the event that one or more users make a decoding error, for a given code (3).

E. Optimal source-channel rate κ^*

A source-channel rate of κ channel symbols per source symbol is said to be *achievable* if the following holds: For any $\epsilon > 0$ there exists non-negative integers m and n (chosen sufficiently large depending on ϵ) together with an (m, n) -code (3) such that

$$\kappa = \frac{n}{m} \quad \text{and} \quad P_e \leq \epsilon.$$

Definition 1: The optimal source-channel rate is

$$\kappa^* := \inf \{ \kappa \geq 0 : \kappa \text{ is achievable} \}.$$

We now present a lower and an upper bounds on κ^* . For any subset $\mathcal{S} \subseteq [1, L]$, let $W_{\mathcal{S}} := (W_\ell : \ell \in \mathcal{S})$ denote the tuple of source variables with indices in \mathcal{S} . Let

$$C_\ell := \log |\mathcal{F}| - \max \{ H(Z), H(N_\ell) \}. \quad (4)$$

Theorem 1:

$$\kappa^* \geq \Psi,$$

where

$$\Psi := \max_{\ell \in [1, L]} \frac{1}{C_\ell} H(W_\ell | W_{\ell^c}).$$

Proof: See Appendix A. ■

Let \mathcal{P} denote the set of all joint probability mass functions (pmfs) on $\mathcal{W}_1 \times \dots \times \mathcal{W}_L$, so that any source (W_1, \dots, W_L) is specified by some $p \in \mathcal{P}$. For brevity, we write $(W_1, \dots, W_L) \sim p$. Let $\mathcal{R}(p)$ denote the set of all non-negative real-valued tuples $\mathbf{r} = (r_1, \dots, r_L)$ satisfying

$$\sum_{i \in \mathcal{S}} r_i \geq H(W_{\mathcal{S}} | W_{\mathcal{S}^c}), \quad \forall \mathcal{S} \subset [1, L]. \quad (5)$$

Theorem 2:

$$\kappa^* \leq \min_{\mathbf{r} \in \mathcal{R}(p)} \Upsilon(\mathbf{r}), \quad (6)$$

where the minimum is attained by a tuple \mathbf{r} on the boundary of $\mathcal{R}(p)$ and

$$\Upsilon(\mathbf{r}) := \max_{\ell \in [1, L]} \frac{1}{C_\ell} \sum_{i \in \ell^c} r_i,$$

Proof: See Appendix B. ■

Theorem 2 is proved using standalone source and channel codes, and, in this context, $\mathcal{R}(p)$ represents the achievable rate region of the underlying distributed source coding problem. The reader may recognise that $\mathcal{R}(p)$ is closely related to the

Slepian-Wolf rate region [24, Sec. 15.4.2]. Indeed, the Slepian-Wolf region for $(W_1, \dots, W_L) \sim p$ is given by

$$\left\{ \mathbf{r} \in \mathcal{R}(p) : \sum_{\ell=1}^L r_\ell \geq H(W_{[1,L]}) \right\}. \quad (7)$$

In other words, $\mathcal{R}(p)$ is the Slepian-Wolf rate region without the total sum-rate constraint. Intuitively, the additional sum-rate constraint in (7) does not play a role in $\mathcal{R}(p)$ and Theorem 2 because user ℓ always has its own source data \mathbf{W}_ℓ as side information. The omission of this constraint is an important characteristic of the rate region $\mathcal{R}(p)$ that shapes much of the following discussion.

F. When Theorem 1 meets Theorem 2

Constraints (5) in the definition of $\mathcal{R}(p)$ dictate that for any $\mathbf{r} \in \mathcal{R}(p)$, we must have that $\Upsilon(\mathbf{r}) \geq \Psi$. When this inequality is an equality, we have the following lemma:

Lemma 1: The lower bound in Theorem 1 meets the upper bound in Theorem 2 if and only if there exists a rate-tuple $\mathbf{r} \in \mathcal{R}(p)$ such that $\Upsilon(\mathbf{r}) = \Psi$. Under this condition, $\kappa^* = \Psi$, and source-channel separation (as specifically described in Appendix B) is optimal.

Proof: Recall that for any $\mathbf{r} \in \mathcal{R}(p)$,

$$\Psi \stackrel{\text{a}}{\leq} \kappa^* \stackrel{\text{b}}{\leq} \min_{\mathbf{r}' \in \mathcal{R}(p)} \Upsilon(\mathbf{r}') \leq \Upsilon(\mathbf{r}),$$

where inequalities (a) and (b) follow from Theorems 1 and 2 respectively. It follows immediately that if $\Upsilon(\mathbf{r}) = \Psi$, then the lower bound in Theorem 1 meets the upper bound in Theorem 2, and $\kappa^* = \Psi$.

Conversely, if the lower and upper bounds meet, then the rate tuple $\mathbf{r} \in \mathcal{R}(p)$ that minimises $\Upsilon(\mathbf{r})$ attains the required condition $\Upsilon(\mathbf{r}) = \Psi$. ■

The main contribution of this paper is to establish nontrivial sufficient conditions for which there exists an $\mathbf{r} \in \mathcal{R}(p)$ such that $\Upsilon(\mathbf{r}) = \Psi$.

III. \mathcal{P}^* — A CLASS OF SOURCES FOR LEMMA 1

The existence of an $\mathbf{r} \in \mathcal{R}(p)$ satisfying $\Upsilon(\mathbf{r}) = \Psi$ depends on both the joint pmf p of the source and the entropies of the channel noises in (2). Such an \mathbf{r} can always be found for the following class of sources, irrespective of the particular channel noise entropies:

$$\mathcal{P}^* := \left\{ p' \in \mathcal{P} : \exists \mathbf{r} \in \mathcal{R}(p') \text{ satisfying } \sum_{i \in \ell^c} r_i = H(W_{\ell^c} | W_\ell'), \quad \forall \ell \in [1, L] \right\}. \quad (8)$$

Proposition 1: If $(W_1, \dots, W_L) \sim p \in \mathcal{P}^*$, then there exists an $\mathbf{r} \in \mathcal{R}(p)$ such that $\Upsilon(\mathbf{r}) = \Psi$.

The class \mathcal{P}^* is a useful regularity condition for Lemma 1. Here are two simple examples.

Example 1: If $L = 2$, then $\mathcal{P}^* = \mathcal{P}$.

Example 2: If $(W_1, \dots, W_L) \sim p$ are independent random variables, then $p \in \mathcal{P}^*$.

We now show how one can establish whether an arbitrary source $(W_1, \dots, W_L) \sim p$ belongs to \mathcal{P}^* by checking whether a specific rate tuple satisfies the conditions in (8). To this end, we re-write the L equalities in (8) as a linear system:

$$\mathbf{r} \mathbf{T} = \mathbf{h}(p). \quad (9)$$

Here we are to solve for the rate vector $\mathbf{r} = [r_1 \ r_2 \ \dots \ r_L]$, where \mathbf{T} is the fixed $(L \times L)$ -matrix

$$\mathbf{T} := \begin{bmatrix} 0 & 1 & \dots & 1 \\ 1 & 0 & & 1 \\ \vdots & & \ddots & \\ 1 & 1 & \dots & 0 \end{bmatrix},$$

and

$$\mathbf{h}(p) := [H(W_{1^c} | W_1) \ H(W_{2^c} | W_2) \ \dots \ H(W_{L^c} | W_L)]. \quad (10)$$

The matrix \mathbf{T} has full rank, and we denote the unique solution of (9) by

$$\mathbf{r}^*(p) := \mathbf{h}(p) \mathbf{T}^{-1}, \quad (11)$$

where $\mathbf{r}^*(p) = [r_1^* \ r_2^* \ \dots \ r_L^*]$ and

$$r_\ell^* = \frac{\|\mathbf{h}(p)\|}{L-1} - H(W_{\ell^c} | W_\ell), \quad \forall \ell \in [1, L]. \quad (12)$$

The conclusion from the above discussion is that we can test whether or not $p \in \mathcal{P}^*$ by numerically checking whether $\mathbf{r}^*(p) \in \mathcal{R}(p)$. The next lemma follows immediately.

Lemma 2:

$$\mathcal{P}^* = \{p' \in \mathcal{P} : \mathbf{r}^*(p') \in \mathcal{R}(p')\}.$$

We now use Lemma 2 to give an example of a source in \mathcal{P}^* , and a source that is not in \mathcal{P}^* . We will see in the next section that all “balanced” sources are in \mathcal{P}^* .

Example 3: Let $B_1, B_2, B_3, B_{1,2}, B_{1,3}$ and $B_{2,3}$ be independent and uniformly distributed Bernoulli random variables. Suppose that random variable associated with user one, W_1 , is string of three bits, $B_1, B_{1,2}$, and $B_{1,3}$, i.e.,

$$W_1 := (B_1, B_{1,2}, B_{1,3}).$$

Similarly, let $W_2 := (B_2, B_{1,2}, B_{2,3})$ and $W_3 := (B_3, B_{1,3}, B_{2,3})$. If p is the joint pmf of (W_1, W_2, W_3) , then

$$\mathcal{R}(p) = \left\{ (r_1, r_2, r_3) \in \mathbb{R}^3 : \begin{array}{ll} r_\ell \geq 1, & \forall \ell \\ r_\ell + r_{\ell'} \geq 3, & \forall \ell \neq \ell' \end{array} \right\},$$

$\mathbf{r}^*(p) = (3/2, 3/2, 3/2)$, and therefore $p \in \mathcal{P}^*$.

Example 4: Remove the Bernoulli variables $B_{1,2}$ and $B_{1,3}$ in Example 3 to obtain $(W_1, W_2, W_3) \sim p$ given by $W_1 := B_1$, $W_2 := (B_2, B_{2,3})$ and $W_3 := (B_3, B_{2,3})$. We have

$$\mathcal{R}(p) = \left\{ (r_1, r_2, r_3) \in \mathbb{R}^3 : \begin{array}{ll} r_\ell \geq 1, & \forall \ell \\ r_1 + r_2 \geq 2 \\ r_1 + r_3 \geq 2 \\ r_2 + r_3 \geq 3 \end{array} \right\},$$

$\mathbf{r}^*(p) = (1/2, 3/2, 3/2)$, and therefore $p \notin \mathcal{P}^*$.

Remark 1: If $p \in \mathcal{P}^*$, then Proposition 1 guarantees that the lower bound in Theorem 1 meets the upper bound in Theorem 2. Otherwise (i.e., if $p \notin \mathcal{P}^*$), there is no such guarantee. The upper and the lower bounds can still be tight,

depending on the particular source model and channel noises. The following example describes such a situation.

Example 5: Let $T_1, T_2, T_3, T_{1,2}, T_{2,3}$, and $T_{1,3}$ be independent random variables with entropies $H(T_1) = H(T_2) = H(T_3) = 1$, $H(T_{1,2}) = H(T_{1,3}) = 3$, and $H(T_{2,3}) = 8$. Let $L = 3$, $W_1 = (T_1, T_{1,2}, T_{1,3})$, $W_2 = (T_2, T_{1,2}, T_{2,3})$, $W_3 = (T_3, T_{1,3}, T_{2,3})$, and p be a source pmf that satisfies these conditions. This gives $H(W_2, W_3|W_1) = 10$, $H(W_1, W_2|W_3) = H(W_1, W_3|W_2) = 5$, $r_1^* = 0$, $r_2^* = r_3^* = 5$. Clearly, $\mathbf{r}^*(p) \notin \mathcal{R}(p)$, and thus $p \notin \mathcal{P}^*$. Consider the following two sets of channel parameters:

- 1) $C_\ell = 1$ for all $\ell \in [1, 3]$: This gives $\Psi = 10$. Choosing $\mathbf{r} = (1, 5, 5) \in \mathcal{R}(p)$, we obtain $\Upsilon(\mathbf{r}) = 10$. The bounds in Theorems 1 and 2 meet for these channel parameters.
- 2) $C_1 = 10$ and $C_2 = C_3 = 4$: This gives $\Psi = 1$. Conditions for $\mathcal{R}(p)$ dictate that $r_2 + r_3 \geq 10$ and $r_1 \geq 1$. This implies $\max\{r_1 + r_2, r_1 + r_3\} = r_1 + \max\{r_2, r_3\} \geq r_1 + \frac{r_2 + r_3}{2} \geq 6$, and consequently, $\Upsilon(\mathbf{r}) \geq \max\{\frac{r_2 + r_3}{C_1}, \frac{r_1 + r_3}{C_2}, \frac{r_1 + r_2}{C_3}\} \geq 1.5 > \Psi$. The bounds do not meet for these channel parameters.

IV. BALANCED SOURCES AND THE I -MEASURE

A. Balanced sources

It is sometimes infeasible or undesirable to apply Lemma 1 by numerically testing whether $\mathbf{r}^*(p) \in \mathcal{R}(p)$. For example, suppose that we need to verify that a source-channel rate κ is achievable for every source within some uncountable set (such a situation is described later in Example 8). In such cases, it is helpful to study more general structural properties of \mathcal{P}^* . The next proposition suggests that \mathcal{P}^* is a rather complicated set, and its proof is omitted.

Proposition 2: \mathcal{P}^* is closed for all L , but it is not convex for any $L \geq 3$.

The next proposition was proved by Ong et al. [19]. The proposition determines the optimal source-channel rate κ^* for a special case of three users and “balanced mutual information” sources.

Proposition 3 (Ong et al. [19, Thm. 1]): If we have $L = 3$ users and the discrete memoryless source (W_1, W_2, W_3) satisfies

$$I(W_i; W_j|W_k) \leq I(W_j; W_k|W_i) + I(W_i; W_k|W_j) \quad (13)$$

for all permutations of $i, j, k \in [1, 3]$, then the optimal source-channel rate is given by $\kappa^* = \Psi$.

It is relatively easy to prove¹ Proposition 3 using the ideas in Section III, as shown below:

Proof: From (12), we get

$$\begin{aligned} r_1^* &:= \frac{1}{2} \left(H(W_1, W_2|W_3) + H(W_1, W_3|W_2) \right. \\ &\quad \left. - H(W_2, W_3|W_1) \right), \\ r_2^* &:= \frac{1}{2} \left(H(W_1, W_2|W_3) + H(W_2, W_3|W_1) \right. \\ &\quad \left. - H(W_1, W_3|W_2) \right), \end{aligned}$$

¹Ong et al. [19] gave a direct proof of Proposition 3 using slightly different techniques.

$$r_3^* := \frac{1}{2} \left(H(W_1, W_3|W_2) + H(W_2, W_3|W_1) \right. \\ \left. - H(W_1, W_2|W_3) \right).$$

By construction, we clearly have

$$\begin{aligned} r_1^* + r_2^* &= H(W_1, W_2|W_3), \\ r_1^* + r_3^* &= H(W_1, W_3|W_2), \\ r_2^* + r_3^* &= H(W_2, W_3|W_1). \end{aligned}$$

Moreover, it follows from (13) that

$$\begin{aligned} r_1^* &\geq H(W_1|W_2, W_3), \\ r_2^* &\geq H(W_2|W_1, W_3), \\ r_3^* &\geq H(W_3|W_1, W_2), \end{aligned}$$

and, therefore, $(r_1^*, r_2^*, r_3^*) \in \mathcal{R}(p)$. This implies that $p \in \mathcal{P}^*$, and Proposition 1 gives the desired result. ■

Given the above proof, it is natural to wonder whether one can find a similar “balanced mutual information” condition that works more generally for $L \geq 3$. It turns out that such a generalisation is possible, and we now formalise this idea.

Fix $L \geq 3$ and $(W_1, \dots, W_L) \sim p$. Consider any nonempty subset

$$\mathcal{K} = \{\ell_1, \dots, \ell_k\} \subseteq [1, L]. \quad (14)$$

The *conditional multiple-mutual information*² between the random variables $(W_{\ell_1}, W_{\ell_2}, \dots, W_{\ell_k})$ was defined by Hekstra and Willems [22, Sec. II.D]

$$\begin{aligned} I(W_{\ell_1}; W_{\ell_2}; \dots; W_{\ell_k} | W_{\mathcal{K}^c}) \\ := \sum_{t=1}^k (-1)^{t-1} \sum_{\substack{\mathcal{T} \subseteq \mathcal{K} \\ \text{s.t. } |\mathcal{T}|=t}} H(W_{\mathcal{T}} | W_{\mathcal{K}^c}). \end{aligned}$$

In this paper, it will be convenient to define $I_\emptyset := 0$ and the notation

$$I_{\mathcal{K}} := I(W_{\ell_1}; W_{\ell_2}; \dots; W_{\ell_k} | W_{\mathcal{K}^c})$$

for any nonempty subset (14).

Definition 2: We say that a source $(W_1, \dots, W_L) \sim p$ is *balanced*³ if

$$\bar{\mu}_k \leq \text{gap}_k \mu_k, \quad (15)$$

holds for all $k \in [2, L-1]$, where

$$\bar{\mu}_k := \max_{\substack{\mathcal{S} \subseteq [1, L] \\ \text{s.t. } |\mathcal{S}|=k}} I_{\mathcal{S}},$$

$$\mu_k := \min_{\substack{\mathcal{S} \subseteq [1, L] \\ \text{s.t. } |\mathcal{S}|=k}} I_{\mathcal{S}},$$

and

$$\text{gap}_k := 1 + \frac{1}{k} \left(\frac{L-1}{2L-k-3} \right).$$

²Conditional multiple-mutual information is also called *conditional k -information* [20]–[22].

³The definition of a balanced source here is different from that by Haitner et al. [25, Sec. 3]. Here, we consider a source consisting of multiple “components”, and require that the components $\{W_{\ell}\}$ have “roughly” the same conditional multiple-mutual informations. Haitner et al.’s balance condition is defined for any pmf, and requires that the pmf be “close to uniform most of the time.”

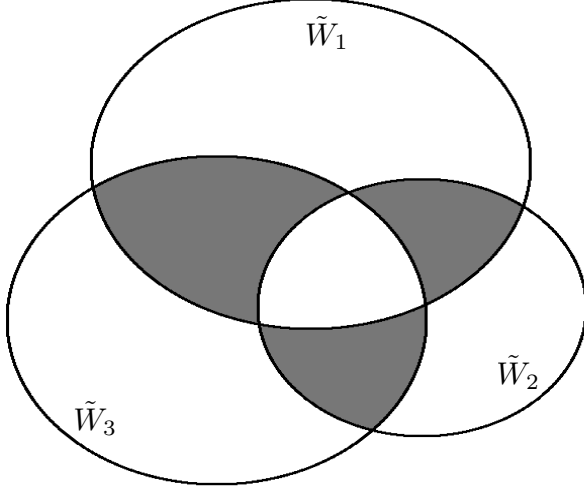


Fig. 2. Conditional multiple-mutual informations I_K , I -measures μ^* , and information diagram for (W_1, W_2, W_3) . The source is balanced if the largest I -measure of the shaded areas is not larger than two times the smallest I -measure of the shaded areas.

Let \mathcal{P}_{bal} denote the set of all balanced sources.

Theorem 3: $\mathcal{P}_{\text{bal}} \subseteq \mathcal{P}^*$.

Proof: See Appendix E. ■

It immediately follows from Theorem 3 that the optimal source-channel rate of any balanced source (regardless of the channel noise entropies) is $\kappa^* = \Psi$. While the set \mathcal{P}_{bal} may not be as large as \mathcal{P}^* , we will see in Section IV-B that \mathcal{P}_{bal} has a measure-theoretic interpretation via the I -measure [23, Chap. 3]. Consequently, checking condition (15) to determine if a source is balanced is equivalent to comparing different areas in *information diagrams*.

The key idea underlying Definition 2 is that a balanced source will have

$$I_K \approx I_{K'}, \quad \forall K, K' \quad \text{with} \quad |K| = |K'|.$$

Here the approximation becomes more stringent as the number of users L and the subset cardinalities grow large, in which case the multiplicative factor gap_k in (15) approaches unity from above.⁴

Condition (15) for balanced sources suggests certain “symmetry” of the source pmf. In particular, if the source pmf p is *symmetrical* in the sense that

$$p(w_1, w_2, \dots, w_L) = p(w_{\ell_1}, w_{\ell_2}, \dots, w_{\ell_L})$$

for all permutations of $\ell_1, \ell_2, \dots, \ell_L \in [1, L]$, then $I_K = I_{K'}$ for all K and K' with $|K| = |K'|$. So, a source with a symmetrical pmf is balanced. We extend this idea in the following example:

Example 6: Consider a random event $B \in \{0, 1\}$ with $\mathbb{P}\{B = 0\} = \rho$, and three sensors each taking a noisy measurement of the event, $W_\ell = B \oplus E_\ell$ for $\ell \in [1, 3]$. Here, $E_\ell \in \{0, 1\}$ is the measurement error with $\mathbb{P}\{E_\ell = 0\} = \sigma_\ell$. If the pmf is symmetrical, i.e., $\sigma_1 = \sigma_2 = \sigma_3$, then the source (W_1, W_2, W_3) is balanced. In addition, since the balance condition (15) does not require all $\{I_K : |K| = k\}$ to be

$$\begin{aligned} I_{\{1\}} &= \mu^*(\tilde{W}_1 \setminus \tilde{W}_{\{2,3\}}) \\ &= \mu^*(\tilde{W}_{\{1,2,3\}}) - \mu^*(\tilde{W}_{\{2,3\}}) \\ &= H(W_{\{1,2,3\}}) - H(W_{\{2,3\}}) \\ &= H(W_{\{1\}}|W_{\{2,3\}}) \\ I_{\{2\}} &= H(W_2|W_{\{1,3\}}) \\ I_{\{3\}} &= H(W_3|W_{\{1,2\}}) \end{aligned}$$

$$\begin{aligned} I_{\{1,2\}} &= \mu^*(\tilde{W}_{\{1\}} \cap \tilde{W}_{\{2\}} \setminus \tilde{W}_{\{3\}}) \\ &= \mu^*(\tilde{W}_{\{1,3\}}) + \mu^*(\tilde{W}_{\{2,3\}}) - \mu^*(\tilde{W}_{\{1,2,3\}}) - \mu^*(\tilde{W}_{\{3\}}) \\ &= H(W_{\{1,3\}}) + H(W_{\{2,3\}}) - H(W_{\{1,2,3\}}) - H(W_{\{3\}}) \\ &= I(W_1; W_2|W_3) \\ I_{\{1,3\}} &= I(W_1; W_3|W_2) \\ I_{\{2,3\}} &= I(W_2; W_3|W_1) \\ I_{\{1,2,3\}} &= \mu^*(\tilde{W}_{\{1\}} \cap \tilde{W}_{\{2\}} \cap \tilde{W}_{\{3\}}) \\ &= I(W_1; W_2) - I(W_1; W_2|W_3) \end{aligned}$$

equal, the source is still balanced if σ_ℓ 's are close, e.g., (a) $\rho = 0.2, \sigma_1 = 0.10, \sigma_2 = 0.12, \sigma_3 = 0.14$; and (b) $\rho = 0.2, \sigma_1 = 0.40, \sigma_2 = 0.41, \sigma_3 = 0.42$. Otherwise, the source is not balanced, e.g., if $\rho = 0.2, \sigma_1 = 0.1, \sigma_2 = 0.12, \sigma_3 = 0.2$.

Balanced conditional mutual-informations lead to balanced conditional entropies in (5) by invoking the next lemma. This lemma plays a key role in the proof of Theorem 3.

Lemma 3:

$$H(W_S|W_{S^c}) = \sum_{K \subseteq S} I_K, \quad \forall S \subseteq [1, L]. \quad (16)$$

Proof: See Appendix C. ■

B. Visualising balanced sources with the I -measure and information diagrams

Definition 2 and Theorem 3 can be visualised using information diagrams and the I -measure [23, Chap. 3]. Fix $L \geq 3$ and the source $(W_1, \dots, W_L) \sim p$. In the notation and terminology of Yeung [23, Chap. 3], let us associate an arbitrary set \tilde{W}_ℓ to each random variable W_ℓ . The I -measure μ^* (defined shortly) is a signed measure on these sets that is chosen in a specific way so that all of Shannon's information measures for (W_1, W_2, \dots, W_L) can be recovered from set-theoretic operations on $\tilde{W}_1, \tilde{W}_2, \dots, \tilde{W}_L$. More specifically, let \mathcal{F}_n denote the *field*⁵ generated by $\tilde{W}_1, \dots, \tilde{W}_L$. For any $S \subseteq [1, L]$, let

$$\tilde{W}_S := \bigcup_{\ell \in S} \tilde{W}_\ell$$

denote the union of all sets with indices in S . The I -measure μ^* on \mathcal{F}_n is defined by

$$\mu^*(\tilde{W}_S) := H(W_S), \quad \text{for all non-empty } S \subseteq [1, L].$$

⁴As a result, we expect the class of \mathcal{P}_{bal} to be relatively smaller as L increases.

⁵The collection of all sets that can be generated from $\tilde{W}_1, \tilde{W}_2, \dots, \tilde{W}_L$ by applying any sequence of the usual set-theoretic operations, i.e., union, intersection, complement, and difference.

It turns out that this signed measure is the only measure that agrees with all Shannon's information measures [23, Thm. 3.9]. For example, the I -measure relates to the mutual information $I(W_1; W_2)$ by

$$\begin{aligned} I(W_1; W_2) &= H(W_1) + H(W_2) - H(W_1, W_2) \\ &= \mu^*(\tilde{W}_1) + \mu^*(\tilde{W}_2) - \mu^*(\tilde{W}_1 \cup \tilde{W}_2) \\ &= \mu^*(\tilde{W}_1 \cap \tilde{W}_2). \end{aligned}$$

Or, more generally, the I -measure μ^* relates to conditional mutual-information via

$$I_K = \mu^* \left(\left(\bigcap_{\ell \in K} \tilde{W}_\ell \right) \setminus \tilde{W}_{K^c} \right).$$

The next example uses μ^* and information diagrams to visualise balanced sources.

Example 7: Consider $L = 3$ users and an arbitrary source (W_1, W_2, W_3) . Figure 2 depicts the corresponding information diagram, and it lists the values of μ^* and conditional mutual-information for all subsets of $\{1, 2, 3\}$. Definition 2 concerns the I -measures of

$$(\tilde{W}_1 \cap \tilde{W}_2) \setminus \tilde{W}_3, \quad (\tilde{W}_1 \cap \tilde{W}_3) \setminus \tilde{W}_2, \quad \text{and} \quad (\tilde{W}_2 \cap \tilde{W}_3) \setminus \tilde{W}_1,$$

which are shaded in Figure 2. In particular, we have

$$\bar{\mu}_2 = \max \{I(W_1; W_2|W_3), I(W_1; W_3|W_2), I(W_2; W_3|W_1)\},$$

$$\underline{\mu}_2 = \min \{I(W_1; W_2|W_3), I(W_1; W_3|W_2), I(W_2; W_3|W_1)\},$$

and the source is balanced if $\bar{\mu}_2 \leq 2\underline{\mu}_2$.

C. Source-channel rate κ and an achievable rate region

The next example shows how one can use Lemma 1 and Theorem 3 to obtain *achievable rates* for the multiway relay channel with common messages.

Example 8: Suppose that we have $L = 3$ users. For each nonempty $\mathcal{S} \subset [1, 3]$, let

$$B_{\mathcal{S}} \in [1, 2^{R_{\mathcal{S}}}]$$

be an independent and uniformly-distributed random variable, where $R_{\mathcal{S}} \geq 0$ and $2^{R_{\mathcal{S}}}$ is an integer. Fix the source-channel rate $\kappa > 0$, and let $(W_1, W_2, W_3) \sim p$ be given by

$$W_1 := (B_{\{1\}}, B_{\{1,2\}}, B_{\{1,3\}}),$$

$$W_2 := (B_{\{2\}}, B_{\{1,2\}}, B_{\{2,3\}}),$$

$$W_3 := (B_{\{3\}}, B_{\{1,3\}}, B_{\{2,3\}}).$$

We wish to characterise the set of all tuples $(R_{\mathcal{S}} : \mathcal{S} \subset [1, 3])$ for which κ is achievable.

We have $I_{\mathcal{S}} = R_{\mathcal{S}}$ for all $\mathcal{S} \subset [1, 3]$. We say that the rate tuple $(R_{\mathcal{S}} : \mathcal{S} \subset [1, 3])$ is balanced if the corresponding source is balanced, that is, when

$$\frac{\max \{R_{\{1,2\}}, R_{\{1,3\}}, R_{\{2,3\}}\}}{\min \{R_{\{1,2\}}, R_{\{1,3\}}, R_{\{2,3\}}\}} \leq 2. \quad (17)$$

Applying Lemma 1 and Theorem 3, we have that κ is achievable for a balanced rate tuple $(R_{\mathcal{S}} : \mathcal{S} \subset [1, 3])$ if (and only if)⁶

$$\kappa > \max \left\{ \frac{R_{\{2\}} + R_{\{3\}} + R_{\{2,3\}}}{\log |\mathcal{F}| - \max \{H(Z), H(N_1)\}} \right\},$$

⁶Replace the strict inequality $>$ with an inequality for the case of only if.

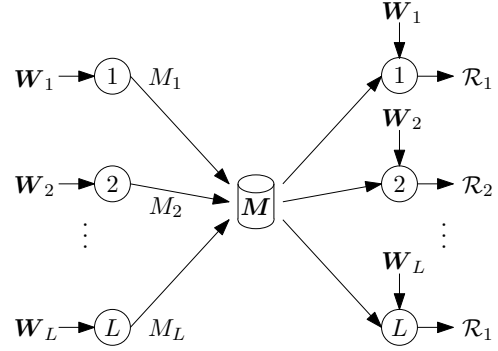


Fig. 3. A centralised storage system with L clients.

$$\left. \begin{aligned} & \frac{R_{\{1\}} + R_{\{3\}} + R_{\{1,3\}}}{\log |\mathcal{F}| - \max \{H(Z), H(N_2)\}}, \\ & \frac{R_{\{1\}} + R_{\{2\}} + R_{\{1,2\}}}{\log |\mathcal{F}| - \max \{H(Z), H(N_3)\}} \end{aligned} \right\}.$$

V. BEYOND RELAYING: AN APPLICATION OF \mathcal{P}^* TO CENTRALISED STORAGE SYSTEMS

The ideas in Sections III and IV concern the class of sources \mathcal{P}^* , and they can be applied to any multiterminal problem for which $\mathcal{R}(p)$ is a meaningful rate region. To illustrate this idea, we now present one such example concerning the centralised storage of correlated data.

A. Problem setup

Consider the data-storage system depicted in Figure 3. The L clients have correlated source data that they wish to write to the storage device. Suppose that the clients' data is generated by the source in Section II, and that the data of client ℓ is the iid m -tuple $\mathbf{W}_\ell = (W_{\ell,1}, W_{\ell,2}, \dots, W_{\ell,m})$. We assume that the method of storage must allow *any* client in the future to reliably recover the source data of *any* subset of clients. We also assume that the storage device is "dumb" in the sense that the clients can read and write, but the device itself does not process the stored data.

Client ℓ writes $M_\ell = f_\ell(\mathbf{W}_\ell)$ to the storage device, where $f_\ell : \mathcal{W}_\ell^m \rightarrow \mathcal{M}_\ell$. At some future time, client ℓ will attempt to recover the source data of all other users source data by computing

$$(\hat{\mathbf{W}}_{\ell,1}, \hat{\mathbf{W}}_{\ell,2}, \dots, \hat{\mathbf{W}}_{\ell,L}) := g_\ell(\mathbf{W}_\ell, M_1, M_2, \dots, M_L),$$

where $g_\ell : \mathcal{W}_\ell \times \mathcal{M}_1 \times \dots \times \mathcal{M}_L \rightarrow \mathcal{W}_1^m \times \dots \times \mathcal{W}_L^m$. We call $(f_1, \dots, f_L, g_1, \dots, g_L)$ an m -code. Let

$$P_e := \mathbb{P} \left[\bigcup_{\ell=1}^L \left\{ g_\ell(M_1, \dots, M_L) \neq (\mathbf{W}_1, \dots, \mathbf{W}_L) \right\} \right]$$

denote the code's average probability of error, and let

$$r_\Sigma := \sum_{\ell=1}^L \frac{1}{m} \log |\mathcal{M}_\ell|$$

denote the total *storage rate* needed by the device (that means the device needs at least mr_Σ bits to store the clients' data).

B. Optimal storage rate

We say that a *storage rate* of r_Σ is *achievable* if for any $\epsilon > 0$ there exists a sufficiently large m and m -code where $P_e \leq \epsilon$. The *optimal storage rate* is

$$r_\Sigma^* := \inf \{r_\Sigma \geq 0 : r_\Sigma \text{ is achievable}\}.$$

The next theorem is proved in Appendix I.

Theorem 4: Considering $(W_1, \dots, W_L) \sim p$,

$$r_\Sigma^* = \min_{\mathbf{r} \in \mathcal{R}(p)} \|\mathbf{r}\|. \quad (18)$$

The next corollary specialises Theorem 4 to give a closed-form expression for the optimal storage rate of any source in \mathcal{P}^* .

Corollary 4.1: If $(W_1, \dots, W_L) \sim p$ with $p \in \mathcal{P}^*$, then

$$r_\Sigma^* = \frac{1}{L-1} \|\mathbf{h}(p)\|,$$

where

$$\mathbf{h}(p) = [H(W_{1^c}|W_1) \ H(W_{2^c}|W_2) \ \dots \ H(W_{L^c}|W_L)].$$

Proof: Recall the rate tuple $\mathbf{r}^*(p) = [r_1^*, \dots, r_L^*]$ in (11) defined by

$$r_\ell^* = \frac{\|\mathbf{h}(p)\|}{L-1} - H(W_{\ell^c}|W_\ell), \quad \forall \ell \in [1, L].$$

Lemma 2 showed that $\mathbf{r}^*(p) \in \mathcal{R}(p)$ if and only if $p \in \mathcal{P}^*$. We will next show that the rate tuple $\mathbf{r}^*(p)$ attains the right hand side of (18), and therefore we have

$$r_\Sigma^* = \|\mathbf{r}^*(p)\|. \quad (19)$$

To see why (19) must be true, let us suppose, to the contrary, that there exists another $\mathbf{r}' \in \mathcal{R}(p)$ with $\|\mathbf{r}'\| < \|\mathbf{r}^*(p)\|$. By construction, $\mathbf{r}^*(p)$ is the unique solution of the linear system

$$\sum_{i \in \ell^c} r_i^* = H(W_{\ell^c}|W_\ell), \quad \forall \ell \in [1, L].$$

Hence, there exists an $\ell \in [1, L]$ such that \mathbf{r}' satisfies

$$\sum_{i \in \ell^c} r'_i < \sum_{i \in \ell^c} r_i^* = H(W_{\ell^c}|W_\ell).$$

This strict inequality leads to the contradiction $\mathbf{r}' \notin \mathcal{R}(p)$, and thus $\mathbf{r}^*(p)$ must achieve the minimum in (18). Finally, (19) simplifies to

$$\begin{aligned} r_\Sigma^* &= \|\mathbf{r}^*(p)\| \\ &= \sum_{\ell=1}^L \left(\frac{\|\mathbf{h}(p)\|}{L-1} - H(W_{\ell^c}|W_\ell) \right) \\ &= \frac{L}{L-1} \|\mathbf{h}(p)\| - \|\mathbf{h}(p)\| \\ &= \frac{\|\mathbf{h}(p)\|}{L-1}. \end{aligned}$$

VI. SUMMARY AND CONCLUSIONS

Finding the optimal source-channel rate κ^* of the multiway relay channel is an open problem whose solution will provide design insights for cooperative communications systems. We presented simple lower and upper bounds on κ^* in Theorems 1 and 2. In Lemma 1, we leveraged these bounds to give a closed-form expression for κ^* and a source-channel separation theorem.

Lemma 1 holds for all combinations of sources and channels where there exists a rate tuple $\mathbf{r} \in \mathcal{R}(p)$ such that $\Upsilon(\mathbf{r}) = \Psi$ (that is, the lower bound in Theorem 1 meets the upper bound in Theorem 2). Here $\mathcal{R}(p)$ is the achievable rate region of the underlying distributed source-coding problem, and the condition $\Upsilon(\mathbf{r}) = \Psi$ depends on both the source and the channel. In general, it remains an open problem to determine κ^* for source-channel combinations where there does not exist such an \mathbf{r} , and for these source-channel combinations, it may be useful to bound the gap between Theorems 1 and 2.

Unfortunately, it can be difficult to determine when Lemma 1 holds, and for this reason we presented two regularity conditions in Sections III and IV. The first regularity condition describes a class of sources \mathcal{P}^* for which Lemma 1 is guaranteed to hold, regardless of the channel. Testing whether or not a given source belongs to \mathcal{P}^* involves solving a linear system (see (11)). The second regularity condition describes a class of balanced sources $\mathcal{P}_{\text{bal}} \subseteq \mathcal{P}^*$ using conditional multiple-mutual informations. This balance condition can be easily understood via the I -measure and information diagrams, and it is most useful in problems where the source is specified by its I -measures (see Example 8).

Finally, the source classes \mathcal{P}^* and \mathcal{P}_{bal} concern only the entropic structure of the distributed source coding rate region $\mathcal{R}(p)$ and, therefore, can be applied to any problem where $\mathcal{R}(p)$ is meaningful. To illustrate this idea, we used \mathcal{P}^* and \mathcal{P}_{bal} to describe an optimal storage rate for a centralised storage problem in Section V.

APPENDIX A PROOF OF THEOREM 1

Suppose that $\kappa > 0$ is achievable ($\kappa = 0$ is trivial). Fix $0 < \epsilon \leq 1/2$. There exists integers m and n with $n/m = \kappa$ and an (m, n) -code (3) satisfying $P_e \leq \epsilon$ for any $\epsilon > 0$. For any $\ell \in [1, L]$,

$$\begin{aligned} mH(W_{\ell^c}|W_\ell) &\stackrel{\text{a}}{=} H(\mathbf{W}_{\ell^c}|\mathbf{W}_\ell), \\ &\stackrel{\text{b}}{\leq} H(\mathbf{W}_{\ell^c}|\mathbf{W}_\ell) - H(\mathbf{W}_{\ell^c}|\mathbf{W}_\ell, \mathbf{Y}_\ell) + \epsilon(m, \epsilon), \\ &\stackrel{\text{c}}{\leq} I(\mathbf{V}; \mathbf{Y}_\ell) + \epsilon(m, \epsilon) \\ &\stackrel{\text{d}}{\leq} \sum_{i=1}^n I(V_i; Y_{\ell,i}) + \epsilon(m, \epsilon) \\ &\stackrel{\text{e}}{\leq} n(\log |\mathcal{F}| - H(N_\ell)) + \epsilon(m, \epsilon). \end{aligned} \quad (20)$$

Notes on (20):

- a. The source is iid.

- b. Noting that $P_e \leq \epsilon \leq 1/2$ and invoking Fano's inequality [24, Thm. 2.10.1], we get $H(\mathbf{W}_{\ell^c} | \mathbf{W}_\ell, \mathbf{Y}_\ell) \leq \varepsilon(m, \epsilon)$, where

$$\varepsilon(m, \epsilon) := 1 + \epsilon m \sum_{i \in \ell^c} \log |\mathcal{W}_i|.$$

- c. By the chain rule and non-negativity of conditional mutual information, we have

$$\begin{aligned} I(\mathbf{W}_{\ell^c}; \mathbf{Y}_\ell | \mathbf{W}_\ell) &\leq I(\mathbf{W}_{\ell^c}, \mathbf{V}; \mathbf{Y}_\ell | \mathbf{W}_\ell) \\ &= I(\mathbf{W}_{\ell^c}, \mathbf{W}_\ell, \mathbf{V}; \mathbf{Y}_\ell) - I(\mathbf{W}_\ell; \mathbf{Y}_\ell) \\ &\leq I(\mathbf{V}; \mathbf{Y}_\ell), \end{aligned}$$

where the last inequality follows because $(\mathbf{W}_{\ell^c}, \mathbf{W}_\ell) \leftrightarrow \mathbf{V} \leftrightarrow \mathbf{Y}_\ell$ forms a Markov chain.

- d. The downlink channel is memoryless; in particular,

$$\begin{aligned} I(\mathbf{V}; \mathbf{Y}_\ell) &= \sum_{i=1}^n I(\mathbf{V}; Y_{\ell,i} | Y_{\ell,1}, \dots, Y_{\ell,i-1}) \\ &\leq \sum_{i=1}^n \left(H(Y_{\ell,i}) - H(Y_{\ell,i} | \mathbf{V}, Y_{\ell,1}^{i-1}) \right) \\ &\stackrel{\text{d.1}}{=} \sum_{i=1}^n \left(H(Y_{\ell,i}) - H(Y_{\ell,i} | V_i) \right), \end{aligned}$$

where (d.1) follows because $Y_{\ell,i} \leftrightarrow V_i \leftrightarrow (V_1^{i-1}, V_{i+1}^n, Y_{\ell,1}^{i-1})$ forms a Markov chain.

- e. $I(V_i; Y_{\ell,i}) = H(Y_{\ell,i}) - H(Y_{\ell,i} | V_i) \leq \log |\mathcal{F}| - H(N_\ell)$, where $H(Y_{\ell,i}) \leq \log |\mathcal{F}|$, and $H(Y_{\ell,i} | V_i) = H(N_\ell)$ follows from the additive-noise channel law $Y_{\ell,i} = V_i \oplus N_\ell$.

By similar arguments, we have

$$mH(W_{\ell^c} | W_\ell) \leq n(\log |\mathcal{F}| - H(Z)) + \varepsilon(m, \epsilon), \quad (21)$$

Combining $\kappa = n/m$ together with (20) and (21), we get

$$\kappa \geq \frac{H(W_{\ell^c} | W_\ell) - \varepsilon(m, \epsilon)/m}{\log |\mathcal{F}| - \max\{H(Z), H(N_\ell)\}}, \quad \forall \ell \in [1, L]. \quad (22)$$

For any $\epsilon > 0$, since (22) must hold for all sufficiently large m , we must have

$$\kappa \geq \frac{H(W_{\ell^c} | W_\ell)}{\log |\mathcal{F}| - \max\{H(Z), H(N_\ell)\}}, \quad \forall \ell \in [1, L]. \quad \blacksquare$$

APPENDIX B PROOF OF THEOREM 2

We use the standard (strict sense) separate source-channel coding technique to prove Theorem 2. The channel capacity and source-coding regions of interest are defined next.

A. Channel capacity region

For each $\ell \in [1, L]$, let $M_\ell \in \mathcal{M}_\ell$ be an independent and uniformly distributed random variable (a channel-coding message) on a finite set \mathcal{M}_ℓ . Recast the joint source-channel coding problem in Section II as a pure channel coding problem with M_ℓ in place of \mathbf{W}_ℓ . More specifically, we

- define an n -code via (3) by setting $m = 1$ and replacing \mathbf{W}_ℓ with M_ℓ and \mathcal{W}_ℓ with \mathcal{M}_ℓ throughout Section II-D, and

- require each user to exchange its message with that of every other user.

For any given n -code, let

$$P_e := \mathbb{P} \left[\bigcup_{\ell=1}^L \left\{ g_\ell(M_\ell, \mathbf{Y}_\ell) \neq (M_1, M_2, \dots, M_L) \right\} \right]$$

denote the average probability of error, and let $\mathbf{R} = (R_1, R_2, \dots, R_L)$ with

$$R_\ell := \frac{1}{n} \log_2 |\mathcal{M}_\ell|$$

denote the channel-coding rates of each user (in bits per channel use).

A channel-coding rate tuple \mathbf{R} is *achievable* if for any $\epsilon > 0$ there exists an n -code such that $P_e \leq \epsilon$. The *capacity region* \mathcal{C} is the closure of set of all achievable rate tuples.

Lemma 4 (Ong et al. [2]):

$$\mathcal{C} = \left\{ \mathbf{R} \in [0, \infty)^L : \sum_{i \in \ell^c} R_i \leq C_\ell, \forall \ell \in [1, L] \right\},$$

where C_ℓ is defined in (4).

B. Source coding region

Consider an arbitrary source $(W_1, \dots, W_L) \sim p$ and recall the setup of Section II-B. Suppose that the users are required to exchange their source data via rate-limited noiseless channels, instead of the noisy finite-field channel. In particular, suppose that user ℓ compresses its source data \mathbf{W}_ℓ to a discrete index $M_\ell := f_\ell(\mathbf{W}_\ell)$, where $f_\ell : \mathcal{W}_\ell^m \rightarrow \mathcal{M}_\ell$. User ℓ is given every index and it attempts to reconstruct the source data of all users:

$$(\hat{\mathbf{W}}_{\ell,1}, \hat{\mathbf{W}}_{\ell,2}, \dots, \hat{\mathbf{W}}_{\ell,L}) := g_\ell(\mathbf{W}_\ell, M_1, M_2, \dots, M_L),$$

where $g_\ell : \mathcal{W}_\ell^m \times \mathcal{M}_1 \times \dots \times \mathcal{M}_L \rightarrow \mathcal{W}_1^m \times \dots \times \mathcal{W}_L^m$. We call the above collection of compressors and decompressors an m -code. For any given m -code, let

$$P_e := \mathbb{P} \left[\bigcup_{\ell=1}^L \left\{ g_\ell(\mathbf{W}_\ell, M_1, \dots, M_L) \neq (\mathbf{W}_1, \dots, \mathbf{W}_L) \right\} \right]$$

denote the average probability of error, and let $\mathbf{r} = (r_1, r_2, \dots, r_L)$ with

$$r_\ell := \frac{1}{m} \log_2 |\mathcal{M}_\ell|$$

denote the source-coding rates of each user (in bits per source symbol).

A source-coding rate \mathbf{r} is *achievable* if for any $\epsilon > 0$ there exists an m -code such that $P_e \leq \epsilon$. The *source coding region* is the closure of the set of all achievable rate tuples.

Lemma 5 (Timo et al. [7]): Let $(W_1, \dots, W_L) \sim p$. The source coding region is equal to $\mathcal{R}(p)$.

C. Source-channel coding with standalone codes

Let us now return to the joint source-channel coding problem. Denote the interiors of \mathcal{C} and $\mathcal{R}(p)$ respectively by

$$\text{int}(\mathcal{C}) := \{\mathbf{a} \in \mathcal{C} : \exists \epsilon > 0 \text{ with } \mathbf{a} + \mathcal{B}_{\mathbf{a}}(\epsilon) \subset \mathcal{C}\}$$

and

$$\text{int}(\mathcal{R}(p)) := \{\mathbf{b} \in \mathcal{R}(p) : \exists \epsilon > 0 \text{ with } \mathbf{b} + \mathcal{B}_{\mathbf{b}}(\epsilon) \subset \mathcal{R}(p)\},$$

where $\mathcal{B}_{\mathbf{a}}(\epsilon) := \{\mathbf{b} \in \mathbb{R}^L : \|\mathbf{a} - \mathbf{b}\| \leq \epsilon\}$. We now prove the first assertion of Theorem 2. Map the output of user ℓ 's source encoder, i.e., $M_\ell \in [1, 2^{mr_\ell}]$, to the input of its channel encoder. This mapping is bijective if and only if $R_\ell = r_\ell/\kappa$. If

$$\mathbf{R} = \mathbf{r}/\kappa \in \text{int}(\mathcal{C}) \quad \text{and} \quad \mathbf{r} \in \text{int}(\mathcal{R}(p)), \quad (23)$$

then each user can separately perform source and channel decoding, to reliably decode its required message. This means the source-channel rate κ is achievable.

We now show that if

$$\kappa > \min_{\mathbf{r} \in \mathcal{R}(p)} \max_{\ell \in [1, L]} \frac{1}{C_\ell} \sum_{i \in \ell^c} r_i, \quad (24)$$

then there exists a rate tuple \mathbf{r} such that (23) holds, and therefore κ is achievable.

Firstly, let $\mathbf{r}^\dagger = (r_1^\dagger, r_2^\dagger, \dots, r_L^\dagger)$ be a rate tuple that attains $\min_{\mathbf{r} \in \mathcal{R}(p)} \max_{\ell \in [1, L]} \frac{1}{C_\ell} \sum_{i \in \ell^c} r_i$. This means the chosen κ in (24) can be written as

$$\kappa = \delta + \max_{\ell \in [1, L]} \frac{1}{C_\ell} \sum_{i \in \ell^c} r_i^\dagger \quad (25a)$$

$$> \max_{\ell \in [1, L]} \frac{1}{C_\ell} \sum_{i \in \ell^c} (r_i^\dagger + \rho) \quad (25b)$$

$$= \max_{\ell \in [1, L]} \frac{1}{C_\ell} \sum_{i \in \ell^c} r'_i, \quad (25c)$$

for some $\delta > 0$, where $\rho := \frac{\delta \min_{\ell \in [1, L]} C_\ell}{L} > 0$, and $r'_i := r_i^\dagger + \rho$.

Now, let $\mathbf{r}' = (r'_1, r'_2, \dots, r'_L)$. Clearly, since $\mathbf{r}^\dagger \in \mathcal{R}(p)$, we have $\mathbf{r}' \in \text{int}(\mathcal{R}(p))$. Also, for each $\ell \in [1, L]$, we select

$$R'_\ell = \frac{r'_\ell}{\kappa} < \frac{r'_\ell}{\max_{k \in [1, L]} \frac{1}{C_k} \sum_{i \in k^c} r'_i}.$$

It follows that, for each $\ell \in [1, L]$,

$$\sum_{j \in \ell^c} R'_j < \frac{\sum_{j \in \ell^c} r'_j}{\max_{k \in [1, L]} \frac{1}{C_k} \sum_{i \in k^c} r'_i} \leq \frac{\sum_{j \in \ell^c} r'_j}{\frac{1}{C_\ell}} \sum_{j \in \ell^c} r'_j = C_\ell.$$

This means $\mathbf{r}'/\kappa \in \text{int}(\mathcal{C})$. Since any κ satisfying (24) is achievable, we have (6).

Finally, since the region $\mathcal{R}(p)$ is closed, and $\max_{k \in [1, L]} \frac{1}{C_k} \sum_{i \in k^c} r_i$ is a strictly-increasing function of any r_i , the right-hand side of (6) is attained by a tuple \mathbf{r} on the boundary of $\mathcal{R}(p)$. This completes the proof of Theorem 2. ■

APPENDIX C PROOF OF LEMMA 3

We now show that

$$H(W_S|W_{S^c}) = \sum_{\mathcal{K} \subseteq \mathcal{S}} I_{\mathcal{K}}, \quad \forall \mathcal{S} \subseteq [1, L]. \quad (26)$$

The proof follows by induction: We first show that (26) holds for all subsets with cardinality 0 and 1. We then show that the truth of (26) for any subset $\mathcal{S} \subset [1, L]$ implies the truth of (26) for all subsets $\mathcal{S}' \subseteq [1, L]$ of cardinality $|\mathcal{S}'| = |\mathcal{S}| + 1$.

Starting with cardinality 0 and the empty set, we have

$$\sum_{\mathcal{K} \subseteq \emptyset} I_{\mathcal{K}} = I_{\emptyset} = H(W_{\emptyset}|W_{[1, L]}) = 0. \quad (27)$$

Now consider any singleton $\{\ell\} \subset [1, L]$. We have

$$\sum_{\mathcal{K} \subseteq \{\ell\}} I_{\mathcal{K}} = I_{\emptyset} + I_{\{\ell\}} = H(W_{\ell}|W_{\ell^c}). \quad (28)$$

Suppose now that we are given $\mathcal{S} \subset [1, L]$ such that

$$\sum_{\mathcal{K} \subseteq \mathcal{S}} I_{\mathcal{K}} = H(W_{\mathcal{S}}|W_{\mathcal{S}^c}). \quad (29)$$

For any $j \in \mathcal{S}^c$, we have

$$\begin{aligned} H(W_{\mathcal{S} \cup \{j\}}|W_{(\mathcal{S} \cup \{j\})^c}) &= H(W_{\mathcal{S} \cup \{j\}}|W_{\mathcal{S}^c \setminus \{j\}}) \\ &\stackrel{a}{=} H(W_{\mathcal{S}}|W_{\mathcal{S}^c \setminus \{j\}}, W_j) \\ &\quad + H(W_j|W_{\mathcal{S}^c \setminus \{j\}}) \\ &\stackrel{b}{=} \sum_{\mathcal{K} \subseteq \mathcal{S}} I_{\mathcal{K}} + H(W_j|W_{\mathcal{S}^c \setminus \{j\}}) \\ &\stackrel{c}{=} \sum_{\mathcal{K} \subseteq \mathcal{S}} I_{\mathcal{K}} + \sum_{\mathcal{K} \subseteq \mathcal{S}} I_{\mathcal{K} \cup \{j\}} \\ &= \sum_{\mathcal{K} \subseteq \mathcal{S} \cup \{j\}} I_{\mathcal{K}}, \end{aligned} \quad (30)$$

where step (a) applies the chain rule for entropy, and step (b) follows by the inductive assumption (29). Step (c) is the key ingredient of our argument, and we prove it separately.

Assuming that step (c) holds, we may now conclude the following: The hypothesis (26) is true for the empty set and all singletons $\{\ell\} \subset [1, L]$ by (27) and (28) respectively. The inductive step (29) holds for all $j \in \mathcal{S}^c$, and hence the hypothesis (26) is true for any subset \mathcal{S} with any cardinality $|\mathcal{S}| \in [2, L]$. The next lemma completes the proof by verifying step (c).

Lemma 6: Let $\mathcal{S} \subset [1, L]$ and $j \in \mathcal{S}^c$ be arbitrary. Then,

$$H(W_j|W_{\mathcal{S}^c \setminus \{j\}}) = \sum_{\mathcal{K} \subseteq \mathcal{S}} I_{\mathcal{K} \cup \{j\}}. \quad (31)$$

Proof: See Appendix D. ■

APPENDIX D PROOF OF LEMMA 6

A. A preliminary lemma

Lemma 7: Let $\mathcal{S} \subset [1, L]$ with $|\mathcal{S}| \leq L - 2$ and $j, m \in \mathcal{S}^c$ with $j \neq m$ be arbitrary. We have

$$I(W_j; W_m|W_{\mathcal{S}^c \setminus \{j, m\}}) = \sum_{\mathcal{K} \subseteq \mathcal{S}} I_{\mathcal{K} \cup \{j, m\}}. \quad (32)$$

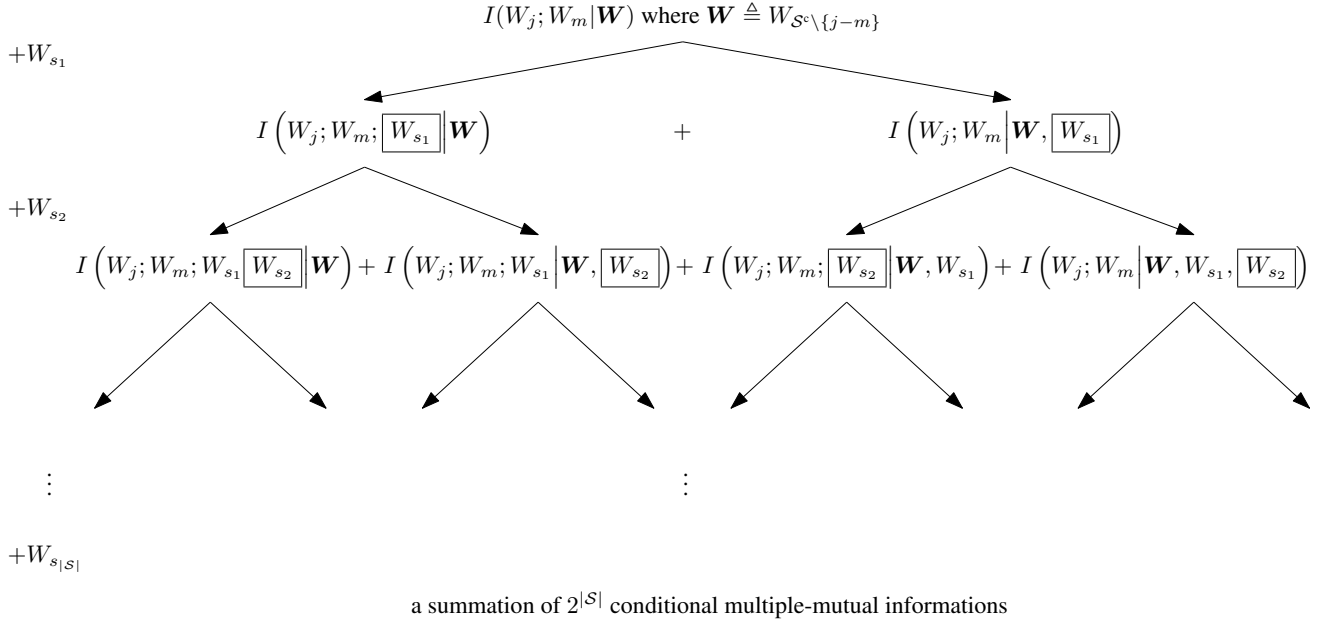


Fig. 4. Recursive method of including each element of $W_S = (W_{s_1}, W_{s_2}, \dots, W_{s_{|S|}})$ into $I(W_j; W_m | W)$

Proof: We first recall a useful identity by Hekstra and Willems [22, Eqn. (10b)]:

$$\begin{aligned} I(W_{\ell_1}; W_{\ell_2}; \dots; W_{\ell_k} | W_{\mathcal{T}}) \\ = I(W_{\ell_1}; W_{\ell_2}; \dots; W_{\ell_k}; W_{\ell_{k+1}} | W_{\mathcal{T}}) \\ + I(W_{\ell_1}; W_{\ell_2}; \dots; W_{\ell_k} | W_{\mathcal{T}}, W_{\ell_{k+1}}), \end{aligned} \quad (33)$$

where $\mathcal{K} = \{\ell_1, \ell_2, \dots, \ell_{k+1}\} \subset [1, L]$ and $\mathcal{T} \subseteq [1, L] \setminus \mathcal{K}$ are arbitrary.

Consider any subset $\mathcal{S} = \{s_1, s_2, \dots, s_{|S|}\} \subset [1, L]$ with $|\mathcal{S}| \leq L - 2$, and some $j, m \in \mathcal{S}^c$ with $j \neq m$. We now take $I(W_j; W_m | W_{\mathcal{S}^c \setminus \{j, m\}})$ and recursively include each element of \mathcal{S} , starting with s_1 , using Hekstra and Willems' identity (33). The procedure is depicted in Figure 4, and it concludes with an expansion consisting of $2^{|S|}$ conditional-multiple-mutual-information terms:

$$\begin{aligned} I(W_j; W_m | W_{\mathcal{S}^c \setminus \{j, m\}}) \\ = \sum_{\mathcal{K} \subseteq \mathcal{S}} I(W_j; W_m; W_{k_1}; \dots; W_{k_{|\mathcal{K}|}} | W_{\mathcal{S}^c \setminus \{j, m\}}, W_{\mathcal{S} \setminus \mathcal{K}}), \end{aligned} \quad (34)$$

where the sum on the right hand side is taken over all subsets of the form $\mathcal{K} = \{k_1, k_2, \dots, k_{|\mathcal{K}|}\} \subseteq \mathcal{S}$. Since $\mathcal{K} \subseteq \mathcal{S}$ and $j, m \in \mathcal{S}^c$, we have

$$\begin{aligned} (\mathcal{S}^c \setminus \{j, m\}) \cup (\mathcal{S} \setminus \mathcal{K}) &= ((\mathcal{S}^c \setminus \mathcal{K}) \cup (\mathcal{S} \setminus \mathcal{K})) \setminus \{j, m\} \\ &= [1, L] \setminus (\mathcal{K} \cup \{j, m\}), \end{aligned}$$

and (34) simplifies to

$$\begin{aligned} I(W_j; W_m | W_{\mathcal{S}^c \setminus \{j, m\}}) \\ = \sum_{\mathcal{K} \subseteq \mathcal{S}} I(W_j; W_m; W_{k_1}; \dots; W_{k_{|\mathcal{K}|}} | W_{[1:L] \setminus (\mathcal{K} \cup \{j, m\})}) \\ = \sum_{\mathcal{K} \subseteq \mathcal{S}} I_{\mathcal{K} \cup \{j, m\}}. \end{aligned} \quad \blacksquare$$

B. Proof of Lemma 6

As before, we use induction to prove (31): We first show that (31) holds for all subsets \mathcal{S} with cardinality 0 and 1. We then show that the truth of (31) for any subset $\mathcal{S} \subset [1, L]$ implies the truth of (31) for all subsets $\mathcal{S}' \subseteq [1, L]$ of cardinality $|\mathcal{S}'| = |\mathcal{S}| + 1$.

Starting with the empty set, $\mathcal{S} = \emptyset$, we have

$$H(W_j | W_{\mathcal{S}^c \setminus \{j\}}) = H(W_j | W_{\{j\}^c}) = \sum_{\mathcal{K} \subseteq \emptyset} I_{\mathcal{K} \cup \{j\}} = I_{\{j\}}. \quad (35)$$

Now consider any $\ell \in [1, L]$ and $\mathcal{S} = \{\ell\}$. We have

$$\begin{aligned} H(W_j | W_{\mathcal{S}^c \setminus \{j\}}) &= H(W_j | W_{[1,L] \setminus \{\ell, j\}}) \\ &= H(W_j | W_{\{j\}^c}) + I(W_j; W_{\ell} | W_{[1,L] \setminus \{\ell, j\}}) \\ &= I_{\{j\}} + I_{\{\ell, j\}} \\ &= \sum_{\mathcal{K} \subseteq \{\ell\}} I_{\mathcal{K} \cup \{j\}}. \end{aligned}$$

Suppose now that we are given $\mathcal{S} \subset [1, L]$ with $|\mathcal{S}| \leq L - 2$ such that (31) holds, i.e.,

$$H(W_j | W_{\mathcal{S}^c \setminus \{j\}}) = \sum_{\mathcal{K} \subseteq \mathcal{S}} I_{\mathcal{K} \cup \{j\}}. \quad (36)$$

Pick any $\ell \in \mathcal{S}^c$ with $\ell \neq j$. We now prove (31) for the set $\mathcal{S} \cup \{\ell\}$. We have

$$\begin{aligned} H(W_j | W_{(\mathcal{S} \cup \{\ell\})^c \setminus \{j\}}) &= H(W_j | W_{\mathcal{S}^c \setminus \{\ell, j\}}) \\ &\stackrel{a}{=} H(W_j | W_{\mathcal{S}^c \setminus \{j\}}) \\ &\quad + I(W_j; W_{\ell} | W_{\mathcal{S}^c \setminus \{\ell, j\}}) \\ &\stackrel{b}{=} \sum_{\mathcal{K} \subseteq \mathcal{S}} I_{\mathcal{K} \cup \{j\}} + \sum_{\mathcal{K} \subseteq \mathcal{S}} I_{\mathcal{K} \cup \{\ell, j\}} \\ &= \sum_{\mathcal{K} \subseteq \mathcal{S} \cup \{\ell\}} I_{\mathcal{K} \cup \{j\}}, \end{aligned} \quad (37)$$

where (a) uses the chain rule for entropy and (b) applies (36) and Lemma 7 and $j \in \mathcal{S}^c$.

We may now conclude the following from the above argument: The hypothesis (31) is true for the empty set and all singletons $\mathcal{S} = \{\ell\}$. The inductive step (37) holds for any $\ell \in \mathcal{S}^c \setminus \{j\}$ and, therefore, the hypothesis (31) is true for any set with any cardinality $|\mathcal{S}| \in [2, L-1]$. ■

APPENDIX E PROOF OF THEOREM 3

If $L = 2$, then $\mathcal{P}^* = \mathcal{P}$ and the theorem is trivial. Suppose that $L \geq 3$ and $(W_1, \dots, W_L) \sim p \in \mathcal{P}_{\text{bal}}$. By Lemma 2, we need only prove that $\mathbf{r}^*(p) \in \mathcal{R}(p)$. We start the proof with a useful lemma that represents the rate tuple $\mathbf{r}^*(p)$ as a weighted sum of conditional multiple-mutual informations.

Let $\mathbf{r}^\dagger(p) = (r_1^\dagger, \dots, r_L^\dagger)$ be defined by

$$r_\ell^\dagger := \sum_{\mathcal{K} \subset [1, L]} J_\ell(\mathcal{K}), \quad \ell \in [1, L], \quad (38)$$

where

$$J_\ell(\mathcal{K}) := \begin{cases} \left(\frac{L - |\mathcal{K}|}{L - 1} \right) I_{\mathcal{K}}, & \text{if } \ell \in \mathcal{K} \\ \left(\frac{1 - |\mathcal{K}|}{L - 1} \right) I_{\mathcal{K}}, & \text{otherwise.} \end{cases} \quad (39)$$

Lemma 8: $\mathbf{r}^\dagger(p) = \mathbf{r}^*(p)$ for all $p \in \mathcal{P}$.

Proof: See Appendix F. ■

We now show that $\mathbf{r}^\dagger(p) \in \mathcal{R}(p)$ by arguing that $\mathbf{r}^\dagger(p)$ satisfies all of the inequalities in (5)—the inequalities defining $\mathcal{R}(p)$ —whenever $p \in \mathcal{P}_{\text{bal}}$. We first notice that Lemma 8 implies that

$$\sum_{i \in \ell^c} r_i^\dagger = H(W_{\ell^c} | W_\ell), \quad \forall \ell \in [1, L].$$

Thus, we need only check the inequality in (5) for all $\mathcal{S} \subset [1, L]$ with cardinality $|\mathcal{S}| \leq L - 2$.

Let $\mathcal{S} \subset [1, L]$ be arbitrary subset with $|\mathcal{S}| \leq L - 2$. Consider the sum

$$\begin{aligned} \sum_{i \in \mathcal{S}} r_i^\dagger &\stackrel{\text{a}}{=} \sum_{i \in \mathcal{S}} \sum_{\mathcal{K} \subset [1, L]} J_i(\mathcal{K}) \\ &\stackrel{\text{b}}{=} \sum_{k=1}^{L-1} \Gamma_k. \end{aligned} \quad (40)$$

Step (a) follows from (38), and in step (b) we define

$$\Gamma_k := \sum_{\substack{\mathcal{K} \subset [1, L] \\ \text{s.t. } |\mathcal{K}|=k}} \sum_{i \in \mathcal{S}} J_i(\mathcal{K}).$$

The next lemma invokes the balanced source assumption and is a key step in the proof.

Lemma 9: Fix $\mathcal{S} \subset [1, L]$ with $|\mathcal{S}| \leq L - 2$. If $p \in \mathcal{P}_{\text{bal}}$, then

$$\Gamma_k \geq \begin{cases} \sum_{\substack{\mathcal{K} \subseteq \mathcal{S} \\ \text{s.t. } |\mathcal{K}|=k}} I_{\mathcal{K}}, & 1 \leq k \leq |\mathcal{S}|, \\ 0, & |\mathcal{S}| < k \leq L - 1. \end{cases}$$

Proof: See Appendix G. ■

Continuing on from (40), we have

$$\sum_{i \in \mathcal{S}} r_i^\dagger = \sum_{k=1}^{L-1} \Gamma_k \stackrel{\text{a}}{\geq} \sum_{k=1}^{|\mathcal{S}|} \sum_{\substack{\mathcal{K} \subseteq \mathcal{S} \\ \text{s.t. } |\mathcal{K}|=k}} I_{\mathcal{K}} \stackrel{\text{b}}{=} H(W_{\mathcal{S}} | W_{\mathcal{S}^c}),$$

where (a) applies Lemma 9 and (b) applies Lemma 3. ■

APPENDIX F PROOF OF LEMMA 8

Recall that $\mathbf{r}^*(p)$, defined in (11), is the unique solution to

$$\sum_{i \in \ell^c} r_i^* = H(W_{\ell^c} | W_\ell), \quad \forall \ell \in [1, L]. \quad (41)$$

Now fix ℓ and consider the same sum over $i \in \ell^c$, but with $\mathbf{r}^*(p)$ replaced by $\mathbf{r}^\dagger(p)$. We have

$$\begin{aligned} \sum_{i \in \ell^c} r_i^\dagger &\stackrel{\text{a}}{=} \sum_{i \in \ell^c} \sum_{\mathcal{K} \subset [1, L]} J_i(\mathcal{K}) \\ &\stackrel{\text{b}}{=} \sum_{\substack{\mathcal{K} \subset [1, L] \\ \mathcal{K} \ni \ell}} \sum_{i \in \ell^c} J_i(\mathcal{K}) + \sum_{\substack{\mathcal{K} \subset [1, L] \\ \mathcal{K} \not\ni \ell}} \sum_{i \in \ell^c} J_i(\mathcal{K}) \\ &\stackrel{\text{c}}{=} \sum_{\mathcal{K} \subseteq \ell^c} I_{\mathcal{K}} \\ &\stackrel{\text{d}}{=} H(W_{\ell^c} | W_\ell). \end{aligned} \quad (42)$$

Lemma (8) now follows directly from (42) and the uniqueness of \mathbf{r}^* . Notes:

- Substitute r_i^\dagger from (38).
- Split the summation over the strict subsets $\mathcal{K} \subset [1, L]$ into two groups: those subsets \mathcal{K} that own ℓ , and those \mathcal{K} that do not own ℓ .
- Consider the sum over subsets \mathcal{K} that do not own ℓ (the second pair of sums in step (b)): The inner sum over i includes $|\mathcal{K}|$ elements with $i \in \mathcal{K}$ and

$$J_i(\mathcal{K}) = \left(\frac{L - |\mathcal{K}|}{L - 1} \right) I_{\mathcal{K}}.$$

The remaining $(L - 1 - |\mathcal{K}|)$ elements with $i \notin \mathcal{K}$ have

$$J_i(\mathcal{K}) = \left(\frac{1 - |\mathcal{K}|}{L - 1} \right) I_{\mathcal{K}}.$$

This observation leads to the expansion shown in (43), which, in turn, simplifies to

$$\sum_{\substack{\mathcal{K} \subset [1, L] \\ \mathcal{K} \not\ni \ell}} \sum_{i \in \ell^c} J_i(\mathcal{K}) = \sum_{\substack{\mathcal{K} \subset [1, L] \\ \mathcal{K} \not\ni \ell}} I_{\mathcal{K}}.$$

Consider the sum over subsets \mathcal{K} that own ℓ . The inner sum over i includes $|\mathcal{K}| - 1$ elements with $i \in \mathcal{K}$ and $(L - 1 - |\mathcal{K}|)$ elements with $i \notin \mathcal{K}$. In this case, we have

$$\sum_{\substack{\mathcal{K} \subset [1, L] \\ \mathcal{K} \ni \ell}} \sum_{i \in \ell^c} J_i(\mathcal{K}) = 0.$$

- Apply Lemma 3.

$$\sum_{\substack{\mathcal{K} \subset [1, L] \\ \mathcal{K} \not\subseteq \ell}} \sum_{i \in \ell^c} J_i(\mathcal{K}) = \sum_{\substack{\mathcal{K} \subset [1, L] \\ \mathcal{K} \not\subseteq \ell}} \left(|\mathcal{K}| \left(\frac{L - |\mathcal{K}|}{L - 1} \right) I_{\mathcal{K}} + (L - 1 - |\mathcal{K}|) \left(\frac{1 - |\mathcal{K}|}{L - 1} \right) I_{\mathcal{K}} \right) \quad (43)$$

	s_1	s_2	s_3	\dots	$s_{ \mathcal{S} }$
\mathcal{K}_1	$J_{s_1}(\mathcal{K}_1)$	$J_{s_2}(\mathcal{K}_1)$	$J_{s_3}(\mathcal{K}_1)$		$J_{s_{ \mathcal{S} }}(\mathcal{K}_1)$
\mathcal{K}_2	$J_{s_1}(\mathcal{K}_2)$	$J_{s_2}(\mathcal{K}_2)$	$J_{s_3}(\mathcal{K}_2)$		$J_{s_{ \mathcal{S} }}(\mathcal{K}_2)$
\mathcal{K}_3	$J_{s_1}(\mathcal{K}_3)$	$J_{s_2}(\mathcal{K}_3)$	$J_{s_3}(\mathcal{K}_3)$		$J_{s_{ \mathcal{S} }}(\mathcal{K}_3)$
\vdots					
$\mathcal{K}_{\binom{L}{k}}$	$J_{s_1}(\mathcal{K}_{\binom{L}{k}})$	$J_{s_2}(\mathcal{K}_{\binom{L}{k}})$	$J_{s_3}(\mathcal{K}_{\binom{L}{k}})$		$J_{s_{ \mathcal{S} }}(\mathcal{K}_{\binom{L}{k}})$

TABLE I
VISUAL AID FOR THE PROOF OF LEMMA 9.

APPENDIX G PROOF OF LEMMA 9

Let $\mathcal{S} = \{s_1, s_2, \dots, s_{|\mathcal{S}|}\} \subset [1, L]$ be any subset with cardinality $|\mathcal{S}| \leq L - 2$, and let $k \in [1, L - 1]$ be arbitrary. Table I will be a useful visual aid throughout the proof. The table consists of

$$\binom{L}{k} := \frac{L!}{(L - k)!k!}$$

rows and $|\mathcal{S}|$ columns—one row for each subset $\mathcal{K} \subset [1, L]$ with cardinality k and one column for each element of \mathcal{S} . Let $\mathcal{K}_1, \mathcal{K}_2, \dots, \mathcal{K}_{\binom{L}{k}}$ be any ordering (for example, lexicographic) of all the subsets \mathcal{K} with cardinality k , and let \mathcal{K}_i be the label for the i -th row of the table. Assign to the cell (\mathcal{K}_i, s_ℓ) the value $J_{s_\ell}(\mathcal{K}_i)$.

We may rewrite Γ_k as a sum over all cells in Table I,

$$\Gamma_k = \sum_{i=1}^{\binom{L}{k}} \sum_{\ell=1}^{|\mathcal{S}|} J_{s_\ell}(\mathcal{K}_i). \quad (44)$$

Recall that

$$J_{s_\ell}(\mathcal{K}_i) = \begin{cases} \left(\frac{L-k}{L-1} \right) I_{\mathcal{K}_i}, & \text{if } s_\ell \in \mathcal{K}_i \\ -\left(\frac{k-1}{L-1} \right) I_{\mathcal{K}_i}, & \text{if } s_\ell \notin \mathcal{K}_i. \end{cases}$$

Definition 3: We call a cell (\mathcal{K}_i, s_ℓ) *positive* if $s_\ell \in \mathcal{K}_i$ and *negative* otherwise.⁷

Lemma 10: In each and every column in Table I, there are

$$\binom{L-1}{k-1} \quad \text{and} \quad \binom{L-1}{k}$$

positive and negative cells respectively.

Proof: Consider an arbitrary column $s_\ell \in \mathcal{S}$. Recall that there are $\binom{L-1}{k-1}$ ways of selecting $(k-1)$ unordered elements from the set $[1, L] \setminus \{s_\ell\}$. The union of each such selection with $\{s_\ell\}$ forms a subset \mathcal{K}_i such that $|\mathcal{K}| = k$ and $\mathcal{K}_i \ni s_\ell$,

⁷The terms *positive* and *negative* refer to the sign of the coefficient fraction in $J_{s_\ell}(\mathcal{K}_i)$, and not to the sign of $I_{\mathcal{K}_i}$.

so it follows that the column has $\binom{L-1}{k-1}$ positive cells. The remaining

$$\binom{L}{k} - \binom{L-1}{k-1} = \binom{L-1}{k}$$

cells in the column are negative. ■

Lemma 11: Throughout the entire table, there are

$$|\mathcal{S}| \binom{L-1}{k-1} \quad \text{and} \quad |\mathcal{S}| \binom{L-1}{k}$$

positive and negative cells respectively.

Proof: The table has $|\mathcal{S}|$ columns and Lemma 10 holds for every column. ■

We now prove the Lemma 9 individually for each of the following three cases: $k = 1$; $2 \leq k \leq |\mathcal{S}|$; and $|\mathcal{S}| + 1 \leq k \leq L - 1$.

A. *Case: $k = 1$*

We trivially have

$$J_i(\mathcal{K}) = \begin{cases} I_{\mathcal{K}}, & \text{if } \mathcal{K} = \{i\}, \\ 0, & \text{otherwise,} \end{cases}$$

and therefore

$$\Gamma_1 = \sum_{i \in \mathcal{S}} I_{\{i\}} = \sum_{\substack{\mathcal{K} \subseteq \mathcal{S} \\ |\mathcal{K}|=1}} I_{\mathcal{K}}. \quad (45)$$

B. *Case: $2 \leq k \leq |\mathcal{S}|$*

We now show that Γ_k is lower bounded as (46). The next definition and lemma will be useful in explaining the steps leading (46).

Definition 4: We say that row \mathcal{K}_i of the Table I is *active* if $\mathcal{K}_i \subseteq \mathcal{S}$ and *inactive* if $\mathcal{K}_i \not\subseteq \mathcal{S}$.

Lemma 12: In Table I, there are

$$k \binom{|\mathcal{S}|}{k} \quad \text{and} \quad |\mathcal{S}| \binom{L-1}{k-1} - k \binom{|\mathcal{S}|}{k}$$

positive cells in active and inactive rows respectively. Similarly, there are

$$(|\mathcal{S}| - k) \binom{|\mathcal{S}|}{k} \quad \text{and} \quad |\mathcal{S}| \binom{L-1}{k} - (|\mathcal{S}| - k) \binom{|\mathcal{S}|}{k}$$

negative cells in active and inactive rows respectively.

Proof: There are $\binom{|\mathcal{S}|}{k}$ active rows in the table and each active row has k positive cells, so there are $k \binom{|\mathcal{S}|}{k}$ positive cells in active rows. The remaining

$$|\mathcal{S}| \binom{L-1}{k-1} - k \binom{|\mathcal{S}|}{k}$$

active cells (here we have used Lemma 11) belong to inactive rows. Similarly, there are $(|\mathcal{S}| - k) \binom{|\mathcal{S}|}{k}$ negative cells in active rows. The remaining

$$|\mathcal{S}| \binom{L-1}{k} - (|\mathcal{S}| - k) \binom{|\mathcal{S}|}{k}$$

$$\begin{aligned}
\Gamma_k &= \sum_{i=1}^{\binom{L}{k}} \sum_{\ell=1}^{|\mathcal{S}|} J_{s_\ell}(\mathcal{K}_i) \\
&\stackrel{\text{a}}{=} \sum_{i=1}^{\binom{L}{k}} \sum_{\ell=1}^{|\mathcal{S}|} \mathbb{1}\{\mathcal{K}_i \subseteq \mathcal{S}\} J_{s_\ell}(\mathcal{K}_i) + \sum_{i=1}^{\binom{L}{k}} \sum_{\ell=1}^{|\mathcal{S}|} (1 - \mathbb{1}\{\mathcal{K}_i \subseteq \mathcal{S}\}) J_{s_\ell}(\mathcal{K}_i) \\
&\stackrel{\text{b}}{=} \sum_{i=1}^{\binom{L}{k}} \mathbb{1}\{\mathcal{K}_i \subseteq \mathcal{S}\} I_{\mathcal{K}_i} + \sum_{i=1}^{\binom{L}{k}} \mathbb{1}\{\mathcal{K}_i \subseteq \mathcal{S}\} \left(\frac{(L - |\mathcal{S}| - 1)(k - 1)}{L - 1} \right) I_{\mathcal{K}_i} + \underbrace{\sum_{i=1}^{\binom{L}{k}} \sum_{\ell=1}^{|\mathcal{S}|} (1 - \mathbb{1}\{\mathcal{K}_i \subseteq \mathcal{S}\}) J_{s_\ell}(\mathcal{K}_i)}_{\text{inactive rows}} \\
&\stackrel{\text{c}}{\geq} \sum_{i=1}^{\binom{L}{k}} \mathbb{1}\{\mathcal{K}_i \subseteq \mathcal{S}\} I_{\mathcal{K}_i} + \binom{|\mathcal{S}|}{k} \left(\frac{(L - |\mathcal{S}| - 1)(k - 1)}{L - 1} \right) \underline{\mu}_k \\
&\quad + \underbrace{\left(|\mathcal{S}| \binom{L - 1}{k - 1} - k \binom{|\mathcal{S}|}{k} \right) \left(\frac{L - k}{L - 1} \right) \underline{\mu}_k - \left(|\mathcal{S}| \binom{L - 1}{k} - (|\mathcal{S}| - k) \binom{|\mathcal{S}|}{k} \right) \left(\frac{k - 1}{L - 1} \right) \bar{\mu}_k}_{\text{inactive rows}} \\
&\stackrel{\text{d}}{=} \sum_{i=1}^{\binom{L}{k}} \mathbb{1}\{\mathcal{K}_i \subseteq \mathcal{S}\} I_{\mathcal{K}_i} + \left(\frac{k - 1}{L - 1} \right) \eta
\end{aligned} \tag{46}$$

negative cells are in inactive rows.

Notes on (46):

- a. Split the outer sum (over rows in Table I) into active and nonactive rows using

$$\mathbb{1}\{\mathcal{K}_i \subseteq \mathcal{S}\} := \begin{cases} 1 & \text{if } \mathcal{K}_i \subseteq \mathcal{S} \\ 0 & \text{otherwise.} \end{cases}$$

- b. There are k positive cells and $(|\mathcal{S}| - k)$ negative cells in each and every active row. Therefore, for every $\mathcal{K}_i \subseteq \mathcal{S}$,

$$\begin{aligned}
\sum_{\ell=1}^{|\mathcal{S}|} J_{s_\ell}(\mathcal{K}_i) &= k \left(\frac{L - k}{L - 1} \right) I_{\mathcal{K}_i} - (|\mathcal{S}| - k) \left(\frac{k - 1}{L - 1} \right) I_{\mathcal{K}_i} \\
&= \left(1 + \frac{(L - |\mathcal{S}| - 1)(k - 1)}{L - 1} \right) I_{\mathcal{K}_i}.
\end{aligned}$$

- c. Use Lemma 12 to count the number of positive and negative cells in the inactive rows (the rightmost pair of sums in step b), and substitute

$$\begin{aligned}
\underline{\mu}_k &= \min\{I_{\mathcal{K}_1}, I_{\mathcal{K}_2}, \dots, I_{\mathcal{K}_{\binom{L}{k}}}\} \\
\text{and } \bar{\mu}_k &= \max\{I_{\mathcal{K}_1}, I_{\mathcal{K}_2}, \dots, I_{\mathcal{K}_{\binom{L}{k}}}\}.
\end{aligned}$$

- d. Clean up the terms (outside the sum in step c) into

$$\eta := \left(\alpha(k, \mathcal{S}) + \frac{1}{k - 1} \beta(k, \mathcal{S}) \right) \underline{\mu}_k - \alpha(k, \mathcal{S}) \bar{\mu}_k,$$

where

$$\alpha(k, \mathcal{S}) := |\mathcal{S}| \binom{L - 1}{k} - (|\mathcal{S}| - k) \binom{|\mathcal{S}|}{k}$$

and

$$\beta(k, \mathcal{S}) := |\mathcal{S}| \binom{L - 1}{k} - (L - 1) \binom{|\mathcal{S}|}{k}.$$

Consider (46):

$$\Gamma_k \geq \left(\frac{k - 1}{L - 1} \right) \eta + \sum_{\substack{\mathcal{K} \subseteq \mathcal{S} \\ \text{s.t. } |\mathcal{K}| = k}} I_{\mathcal{K}},$$

and, in particular, the constants $\alpha(k, \mathcal{S})$ and $\beta(k, \mathcal{S})$ that make up η . We have $\alpha(k, \mathcal{S}) > 0$ and $\beta(k, \mathcal{S}) > 0$, so it follows that $\eta \geq 0$ whenever

$$\bar{\mu}_k \leq \left(1 + \frac{\beta(k, \mathcal{S})}{(k - 1) \alpha(k, \mathcal{S})} \right) \underline{\mu}_k. \tag{47}$$

The next lemma shows that (47) does indeed hold whenever $p \in \mathcal{P}_{\text{bal}}$, and therefore

$$\Gamma_k \geq \sum_{\substack{\mathcal{K} \subseteq \mathcal{S} \\ \text{s.t. } |\mathcal{K}| = k}} I_{\mathcal{K}}. \tag{48}$$

Lemma 13: Fix $(W_1, \dots, W_L) \sim p$ with $p \in \mathcal{P}_{\text{bal}}$ and $2 \leq k \leq L - 2$. For any subset $\mathcal{T} \subset [1, L]$ with $k \leq |\mathcal{T}| \leq L - 2$, we have (47).

Proof: See Appendix H. ■

C. Case: $|\mathcal{S}| \leq k \leq L - 1$

We have

$$\begin{aligned}
\Gamma_k &= \sum_{i=1}^{\binom{L}{k}} \sum_{\ell=1}^{|\mathcal{S}|} J_{s_\ell}(\mathcal{K}_i) \\
&\stackrel{\text{a}}{\geq} \binom{L - 1}{k - 1} \left(\frac{L - k}{L - 1} \right) \underline{\mu}_k - \binom{L - 1}{k} \left(\frac{k - 1}{L - 1} \right) \bar{\mu}_k \\
&= \frac{(L - 2)!}{(L - k - 1)! k!} (k \underline{\mu}_k - (k - 1) \bar{\mu}_k) \\
&\stackrel{\text{b}}{\geq} 0.
\end{aligned} \tag{49}$$

Notes:

- Use Lemma 11 to count the number of positive and negative cells Table I, and bound the corresponding conditional multiple-mutual informations by $\bar{\mu}_k$ and $\underline{\mu}_k$.
- For all $k \in [2, L]$, we have

$$\underline{\mu}_k \stackrel{\text{b.1}}{\leq} \bar{\mu}_k \stackrel{\text{b.2}}{\leq} \left(1 + \frac{1}{k} \left(\frac{L-1}{2L-k-3}\right)\right) \underline{\mu}_k \stackrel{\text{b.3}}{\leq} \left(\frac{k}{k-1}\right) \underline{\mu}_k.$$

Step (b.1) follows by definition of $\underline{\mu}_k$ and $\bar{\mu}_k$; step (b.2) follows because the source is balanced, $p \in \mathcal{P}_{\text{bal}}$; and step (b.3) follows because

$$1 + \frac{1}{k} \left(\frac{L-1}{2L-k-3}\right) \leq \frac{1}{k-1}, \quad \forall k \in [2, L-1].$$

It follows that $k\underline{\mu}_k - (k-1)\bar{\mu}_k \geq 0$, since $0 \leq \underline{\mu}_k \leq \bar{\mu}_k$.

APPENDIX H PROOF OF LEMMA 13

Fix $2 \leq k \leq L-2$. Let $\mathcal{T} \subset [1, L]$ be any subset with cardinality $k \leq |\mathcal{T}| \leq L-2$. We have

$$\frac{\beta(k, \mathcal{T})}{\alpha(k, \mathcal{T})} \geq \frac{k-1}{k} \left(\frac{L-1}{2L-k-3}\right),$$

and it then follows that $p \in \mathcal{P}_{\text{bal}}$ implies

$$\bar{\mu}_k \leq \left(1 + \frac{\beta(k, \mathcal{T})}{(k-1)\alpha(k, \mathcal{T})}\right) \underline{\mu}_k. \quad \blacksquare$$

APPENDIX I PROOF OF THEOREM 4

The centralised storage problem is equivalent to the distributed source coding problem in Appendix B-B. By Lemma 8, a total storage rate $r_\Sigma \geq 0$ is achievable if and only if there exists a rate tuple $\mathbf{r} \in \text{int}(\mathcal{R}(p))$ such that $r_\Sigma \geq \|\mathbf{r}\|$. The optimal total storage rate r_Σ^* is then

$$r_\Sigma^* = \inf_{\mathbf{r} \in \text{int}(\mathcal{R}(p))} \|\mathbf{r}\| = \min_{\mathbf{r} \in \mathcal{R}(p)} \|\mathbf{r}\|. \quad \blacksquare$$

REFERENCES

- [1] D. Gündüz, A. Yener, A. Goldsmith, and H. V. Poor, "The multiway relay channel," *IEEE Trans. Inf. Theory*, vol. 59, no. 1, pp. 51–63, Jan. 2013.
- [2] L. Ong, S. J. Johnson, and C. M. Kellett, "The capacity region of multiway relay channels over finite fields with full data exchange," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 3016–3031, May 2011.
- [3] J. G. Andrews, W. Choi, and R. W. Heath, "Overcoming interference in spatial multiplexing MIMO cellular networks," *IEEE Trans. Wirel. Commun.*, vol. 14, no. 6, pp. 95–104, Dec. 2007.
- [4] J.-M. Park, D.-S. Oh, and D.-C. Park, "Coexistence of mobile-satellite service system with mobile service system in shared frequency bands," *IEEE Trans. Consumer Electron.*, vol. 55, no. 3, pp. 1051–1055, Aug. 2009.
- [5] S. Katti, H. Rahul, W. Hu, D. Katabi, M. Médard, and J. Crowcroft, "XORs in the air: Practical wireless network coding," *IEEE-ACM Trans. Netw.*, vol. 16, no. 3, pp. 497–510, June 2008.
- [6] T. C. Jephson, "The basics of reliable distributed storage networks," *IT Prof.*, vol. 6, no. 3, pp. 18–24, May–June 2004.
- [7] R. Timo, G. Lechner, L. Ong, and S. J. Johnson, "Multi-way relay networks: Orthogonal uplink, source-channel separation and code design," *IEEE Trans. Commun.*, vol. 61, no. 2, pp. 753–768, Feb. 2013.
- [8] L. Ong and S. J. Johnson, "The capacity region of restricted multi-way relay channels with deterministic uplinks," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Cambridge, USA, July 1–6 2012, pp. 786–790.
- [9] J. Barros and S. Servetto, "Network information flow with correlated sources," *IEEE Trans. Inf. Theory*, vol. 52, no. 1, pp. 155–170, 2006.
- [10] T. Cui, T. Ho, and J. Klierwer, "Memoryless relay strategies for two-way relay channels," *IEEE Trans. Commun.*, vol. 57, no. 10, pp. 3132–3143, 2009.
- [11] T. Cui, J. Klierwer, and T. Ho, "Communication protocols for n-way all-cast relay networks," *IEEE Trans. Commun.*, vol. PP, no. 99, pp. 1–13, 2012.
- [12] B. Rankov and A. Wittneben, "Achievable rate regions for the two-way relay channel," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Seattle, USA, July 9–14 2006, pp. 1668–1672.
- [13] H.-I. Su and A. El Gamal, "Two-way source coding through a relay," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Austin, USA, June 13–18 2010, pp. 176–180.
- [14] A. Wyner, J. Wolf, and F. Willems, "Communicating via a processing broadcast satellite," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1243–1249, 2002.
- [15] L. R. Varshney, J. Kusuma, and V. K. Goyal, "Malleable coding with fixed reuse," *arXiv preprint*, no. arXiv:0809.0737, 2011.
- [16] S. S. Pradhan and K. Ramchandran, "Distributed source coding using syndromes (DISCUS): design and construction," *IEEE Trans. Inf. Theory*, vol. 49, no. 3, pp. 626–643, 2003.
- [17] S. S. Pradhan, J. Kusuma, and K. Ramchandran, "Distributed compression in a dense microprocessor network," *IEEE Signal Process. Mag.*, vol. 19, no. 2, pp. 51–60, 2002.
- [18] C. Y. Wang, S. H. Lim, and M. Gastpar, "Information-theoretic caching: sequential coding for computing," *arXiv preprint*, no. arXiv:1504.00553, 2015.
- [19] L. Ong, G. Lechner, S. J. Johnson, and C. M. Kellett, "The three-user finite-field multi-way relay channel with correlated sources," *IEEE Trans. Commun.*, vol. 61, no. 8, pp. 3125–3135, Aug. 2013.
- [20] W. J. McGill, "Multivariate information transmission," *IRE Trans. Inf. Theory*, vol. 4, no. 4, pp. 93–111, Sept. 1954.
- [21] T. S. Han, "Multiple mutual informations and multiple interactions in frequency data," *Inf. and Control*, vol. 46, no. 1, pp. 26–45, July. 1980.
- [22] A. P. Hekstra and F. M. J. Willems, "Dependence balance bounds for single-output two-way channels," *IEEE Trans. Inf. Theory*, vol. 35, no. 1, pp. 44–53, Jan. 1989.
- [23] R. W. Yeung, *Information Theory and Network Coding*, 1st ed. Springer, 2008.
- [24] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Wiley-Interscience, 2006.
- [25] I. Haitner, O. Horvitz, J. Katz, C.-Y. Koo, R. Morselli, and R. Shaltiel, "Reducing complexity assumptions for statistically-hiding commitment," in *Advances in Cryptology – EUROCRYPT 2005*, R. Cramer, Ed. Springer Berlin Heidelberg, 2005, vol. 3494, pp. 58–77.